Abstracts

Reversible Proposal MCMC with Heavy-tailed Target Distributions

Kengo Kamatani

Osaka University, Japan

ABSTRACT

We will discuss Markov chain Monte Carlo methods for heavy-tailed target probability distributions, based on a reversible proposal transition kernel. We will study the dimensionality effect using the high-dimensional asymptotic analysis of Roberts, Gelman, and Gilks. We also study ergodic properties for heavy-tailed target distributions.

Unbiased inference for discretely observed hidden Markov model diffusions

JORDAN FRANKS[a], AJAY JASRA[b], KODY LAW[c], AND MATTI VIHOLA[a]

[a] *University of Jyväskylä, Finland*
[b] *National University of Singapore, Singapore*
[c] *University of Manchester, United Kingdom*

ABSTRACT

In this talk we discuss an importance sampling (IS) type estimator for Bayesian joint inference on the model parameters and latent states of a class of hidden Markov models (HMMs). We are interested in the class of HMMs for which the hidden state dynamics is a diffusion process and noisy observations are obtained at some number of points in time. We suppose that the diffusion dynamics can not be simulated exactly and hence one must time-discretise the diffusion. Our approach is based on particle marginal Metropolis–Hastings (PMMH), particle filters (PFs), and randomised multilevel Monte Carlo (rMLMC). The estimator is built upon a single run of PMMH using a coarse discretisation of the model. The consequent IS type correction is based on a single-term rMLMC estimator using output from a PF developed for level difference integral estimation. The resulting IS type estimator leads to inference without a bias from the time-discretisation. We give convergence results, such as a central limit theorem, and recommend allocations for algorithm inputs. A nice advantage of our method is its general applicability, as it does not rely on exact simulation of the diffusion. The method is also highly parallelisable, and likely computationally efficient in many situations. We demonstrate the method on two examples from the literature. This talk is based on [1].

References

[1] J. Franks, A. Jasra, K. Law, and M. Vihola. Unbiased inference for discretely observed hidden Markov model diffusions. Preprint arXiv:1807.10259, 2018.

An Overview of Parallel and Distributed MCMC Methods

Cheng Li

National University of Singapore

ABSTRACT

Standard posterior sampling algorithms, such as Markov chain Monte Carlo (MCMC), face major challenges in scaling up to massive datasets with large numbers of observations. In this talk, we focus on the divide-and-conquer (d&c) strategy that has been popularized in the past few years. The d&c Bayesian methods typically consist of three steps: (i) partition the full datasets into many non-overlapping subsets; (ii) run MCMC in parallel on all subsets using stochastic approximations; (iii) combine subset posterior samples into a global approximation to the true posterior distribution. We survey the various d&c Bayesian methods, and discuss their ideas, properties and practical implementation. Our comparisons cover both theoretical analysis and numerical experiments on some commonly used statistical models.

An Easy-to-use Empirical Likelihood ABC Method

SANJAY CHAUDHURI, SHUBROSHEKHAR GHOSH, DAVID NOTT, AND KIM CUC PHAM

National University of Singapore

ABSTRACT

Many scientifically well-motivated statistical models in natural, engineering and environmental sciences are specified through a generative process, but in most cases it may not be possible to write down a likelihood for these models analytically. Approximate Bayesian computation (ABC) methods, which allow Bayesian inference in these situations, are typically computationally intensive. Recently, empirical likelihood based ABC methods, which are computationally attractive, have been suggested in the literature. The current empirical likelihood methods rely on the availability of a set of suitable analytically tractable estimating equations. We propose an easy-to-use empirical likelihood ABC method, where the only inputs required are a choice of summary statistic, it's observed value, and the ability to simulate summary statistics for any parameter value under the model. It is shown that the posterior obtained using the proposed method is consistent, and its performance is explored using various examples.

Multilevel Monte Carlo and Transport Maps

Jeremie Houssineau

National University of Singapore

ABSTRACT

The multilevel Monte Carlo approach has proved useful in applications where solutions of diverse computational cost and accuracy are available; the most straightforward example being given by varying discretizations of SDEs. However, the success of this approach is conditioned on the ability to produce highly-correlated samples between adjacent levels, which can be difficult in many cases. In this talk, I will show how transport maps can be used to address this challenge and I will demonstrate the performance of this approach on SDE-related inference problems as well as on smoothing problems for hidden Markov models.

Bayesian Changepoint Detection in Cyber-security

Nick Heard

Imperial College London, UK

ABSTRACT

Changepoint models have numerous applications within statistical cyber-security: Examples include determining the genealogy of a malware executable; providing flexible models for densities and intensities of normal computer network traffic; and, of course, network anomaly detection to potentially reveal the presence of a network intrusion. This talk will review changepoint inferential tools, and demonstrate how these tools can be applied and adapted in the cyber context.

The Bouncy Particle Sampler in Practice

Alexandre Thiery[a] and Chris Sherlock[b]

[a] *National University of Singapore*
[b] *Lancaster University, UK*

ABSTRACT

The bouncy particle sampler is a promising MCMC algorithm, especially when combined with sub-sampling strategies. In this talk, I will describe a discrete version of it that completely bypass the need for simulating the often complicated "bounce times". Using diffusion approximations and homogenization arguments, I will describe the mixing properties of the algorithm in high-dimensional settings and discuss the tuning of the refreshment rate, the crucial parameter that heavily influences the efficiency of the algorithm.

Hamiltonian Descent Methods for Optimization

Arnaud Doucet, Chris Maddison, Daniel Paulin, and Yee Whye Teh

University of Oxford, UK

ABSTRACT

Hamiltonian dynamics is widely used for sampling. In this talk, we will introduce a new convex optimization method based on Hamiltonian dynamics, and show that it has convergence properties that improve upon existing methods in the literature.

Adaptive Community Detection via Fused l-1 Penalty

YUNJIN CHOI, VINCENT Y.F. TAN, AND ZHAOQIANG LIU

National University of Singapore

ABSTRACT

In recent years, community detection has been an active research area in various fields including machine learning and statistics. While a plethora of works has been published over the past few years, most of the existing methods depend on a predetermined number of communities. Given the situation, determining the proper number of communities is directly related to the performance of these methods. Currently, there does not exist a golden rule for choosing the ideal number, and people usually rely on their background knowledge of the domain to make their choices. To address this issue, we propose a community detection method that also adaptively finds the number of the underlying communities. Central to our method is fused l-1 penalty applied on an induced graph from the given data. The proposed method shows promising results.

Gaussian Variational Approximation with Structured Covariance Matrices

David Nott

National University of Singapore

ABSTRACT

Variational approximation methods provide a scalable alternative to Markov chain Monte Carlo methods in Bayesian computational problems. In this talk we consider approximations to a posterior distribution from a Gaussian family, when the model parameter is high-dimensional. Learning a Gaussian variational approximation is very challenging in high dimensions, because the number of parameters in the variational posterior covariance matrix grows quadratically with the number of model parameters, which makes the variational optimization a very high-dimensional one unless further restrictions are made. In this talk, we consider imposing sparsity in the precision matrix of such an approximation to reflect appropriate conditional independence structure in the model, which allows the Gaussian variational distribution to be both flexible and parsimonious. The sparsity is achieved through parameterization in terms of the Cholesky factor, and efficient stochastic gradient methods are developed for the optimization. Alternative methods for structuring the covariance matrix of the posterior based on factor models will also be discussed briefly, and our approaches will be illustrated using generalized linear mixed models and state-space models for time series.

Spatially smooth local ensemble transform particle filtering

Matthew Graham and Alexandre H. Thiery

National University of Singapore

ABSTRACT

State inference in spatially-extended dynamical systems is a challenging problem with significant practical applications such as numerical weather prediction. Particle filters (PFs) while having attractive theoretical properties require infeasibly large ensemble sizes for accurate inference in high-dimensional spatial models. Localisation approaches, which exploit low dependence between state variables at spatially distant points by performing local state updates, offer a potential resolution to this issue. Naively applying the resampling step of the PF update locally can however produce implausible spatially discontinuous states. The ensemble transform PF replaces resampling with linear transformation by an optimal transport (OT) map and can be localised by computing OT maps for every spatial mesh point. The resulting scheme is however computationally intensive for dense meshes and still produces non-smooth states. In this talk I will present a new local ensemble transform PF method which computes a fixed number of OT maps independent of the mesh resolution and smoothly interpolates these maps across space, reducing the computation required while also ensuring state particles retain spatial smoothness properties. I will illustrate the performance of the proposed approach compared to alternative methods in several nonlinear spatiotemporal models, including a challenging two-dimensional stochastic Navier-Stokes example.

Utilising Inference in State-space Models with Multiple Paths from Conditional Sequetial Monte Carlo

Christophe Andrieu[a], Nicolas Chopin[b], Arnaud Doucet[c], and Sinan Yildirim[d]

[a] *University of Bristol, UK*
[b] *ENSAE Paris Tech, France*
[c] *Oxford University, UK*
[d] *Sabanci University, Turkey*

ABSTRACT

We consider a state-space model $\{X_t, Y_t\}_{t\geq 1}$ with a static parameter θ governing its transition and observation probability laws. Our work concerns Bayesian inference of θ given $Y_{1:n}$ for some $n \geq 1$. When the state-space model is non-linear or non-Gaussian, the inference is utilised with sequential Monte Carlo (SMC). In particular, in the Metropolis-within-particle Gibbs algorithm, an iteration consists of (i) updating the sample for $X_{1:n}$ via a conditional SMC (cSMC), which is followed by (ii) a Metrpolis-Hastings update for θ [1]. Retaining one path from the samples in the cSMC involved in Metropolis-within- particle Gibbs may seem to be wasteful. A natural question is whether it is possible to make use of multiple, even all possible, trajectories and average out the corresponding acceptance ratios. We show that this is possible via the use of asymmetric acceptance ratios. The proposed schemes reduce asymptotic variance at a cost that can be parallellised. This a part of the work in [2] where the use of asymmetric acceptance ratios in Metropolis-Hasting algorithm is studied in more generally.

References

[1] C. Andrieu, A. Doucet, R. Holenstein, *Particle Markov chain Monte Carlo methods*, Journal of the Royal Statistical Society: Series B (Statistical Methodology), 2010, vol 72, issue 3, pp. 269-342.

[2] C. Andrieu, A. Doucet, S. Yildirim, N. Chopin, *On the utility of Metropolis-Hastings with asymmetric acceptance ratio*, arXiv:1803.09527, April, 2018.

Bayesian Inference for Multiple Gaussian Graphical Models with Application to Metabolic Association Networks

Linda Tan

National University of Singapore

ABSTRACT

We investigate the effect of cadmium (an environmental pollutant) on the correlation structure of a number of urinary metabolites using Gaussian graphical models (GGMs). Differential networks, which highlight changes in metabolite interactions under different experimental conditions, are constructed using the fitted GGMs. Analysis of such networks can reveal differences in the underlying biological reactions caused by cadmium exposure. We consider Bayesian inference and propose using the multiplicative (Chung-Lu random graph) model as a prior on the graphical space. In this model, each edge is chosen independently with probability equal to the product of the connectivities of the end nodes. This class of prior is parsimonious yet highly flexible; it can be used to encourage sparsity or graphs with a pre-specified degree distribution when such prior knowledge is available. We extend the multiplicative model to multiple GGMs linking the probability of edge inclusion through logistic regression and demonstrate how this leads to joint inference for multiple GGMs. A sequential Monte Carlo (SMC) algorithm is developed for estimating the posterior distribution of the graphs.

Bayesian Nonparametric Modelling of Recurrent Event Processes

Maria De Iorio[a], Alessandra Guglielmi[b], Giorgio Paulon[c], Marta Tallarita[d], and Francesca Ieva[b]

[a] *University College London, UK and Yale-NUS College, Singapore*
[b] *Politecnico di Milano, Italy*
[c] *The University of Texas at Austin, USA*
[d] *University College London, UK*

ABSTRACT

Heart failure (HF) is one of the main causes of morbidity, hospitalization and death in the western world and the economic burden associated with HF management is relevant and expected to increase in the future. We consider hospitalization data for heart failure in the most populated Italian Region, Lombardia. Data were extracted from the administrative data warehouse of the regional healthcare system. The main clinical outcome of interest is time to death and research focus is on investigating how recurrent hospitalizations affect the time to event. The main contribution of the paper is to develop a joint model for gap times between consecutive re-hospitalizations and survival time. The probability models for the gap times and for the survival outcome share a common patient specific frailty term. Using a flexible Dirichlet process model for the random effects distribution accounts for patient heterogeneity in recurrent event trajectories. Moreover, the joint model allows for dependent censoring of gap times by death or administrative reasons and for the correlations between different gap times for the same individual. It is straightforward to include covariates in the survival and/or recurrence process through the specification of appropriate regression terms. The main advantages of the proposed methodology are wide applicability, ease of interpretation and efficient computations. Posterior inference is implemented through Markov chain Monte Carlo methods.

Localization for MCMC

Xin Tong

National University of Singapore

ABSTRACT

We investigate how ideas from covariance localization in numerical weather prediction can be used to construct effective Markov chain Monte Carlo (MCMC) methods for sampling high-dimensional distributions with banded covariance and precision matrices. The main idea is to exploit banded structure during problem formulation and numerical solution. In particular, we propose to solve high-dimensional Bayesian inverse problems with nearly banded structure (i.e., small off-diagonal elements) by first replacing the problem with a banded version, and then solving the modified problem using a Metropolis-within-Gibbs sampler that exploits this banded structure. We discuss conditions under which posterior moments of the modified problem are close to those of the original problem. Under the same conditions, the convergence rate of an associated sampler is independent of dimension. We present our ideas in the context of Gaussian problems, where mathematical formulations are precise and for which convergence analysis can be made rigorous.

Piecewise Deterministic MCMC

Alexandre Bouchard-Côté[a], George Deligiannidis[b], Arnaud Doucet[b], and Paul Vanetti[b]

[a] *University of British Columbia, Canada*
[b] *University of Oxford, UK*

ABSTRACT

Non-reversible Markov chain Monte Carlo schemes based on piecewise deterministic Markov processes have been recently introduced in applied probability, automatic control, physics and statistics. Although these algorithms demonstrate experimentally good performance and are accordingly increasingly used in a wide range of applications, our theoretical understanding of their properties is still far from complete. I will give an overview of recent progress in this direction.

Asset return predictability from a Bayesian perspective

Benjamin Poignard[a] and Julien Pénasse[b]

[a] *Osaka University, Japan*
[b] *University of Luxembourg, Luxembourg*

ABSTRACT

A distinctive property of long term asset allocation is the time-varying risk-premium, implying the existence of a predictable component among assets. However, an inherent drawback of a long-term analysis is the small number of observations, which hampers any stable predictive relationship. Moreover, the selection of the relevant predictive variables to forecast the risk premium is a tricky challenge. To take both the risk estimation into account - due to the short sample size - and the risk model - due to the imperfect choice of the predictive variables -, we propose to introduce Bayesian priors in the predictive model. This approach offsets the mean-reversion belief and thus impacts the allocation. Furthermore, we model a Bayesian investor who has access to international data. Learning across countries arises whenever this investor believes that international return processes share a common distribution. The proposed model allows for samples of different lengths and introduces financial constraints on equity premium forecasts. Empirically, estimates are more reliable, an effect that manifests itself both in- and out-of-sample. International predictability also appears much less heterogeneous than previously reported.